OPTIMAL DESIGN IN NONLINEAR MULTIRESPONSE ESTIMATION: POISSON MODEL FOR FILTER FEEDING

Kinley Larntz, Dept. of Applied Statistics
Christos Hatzis, Dept. of Chemical Engineering

School of Statistics, University of Minnesota

Technical Report #558

February, 1991

Optimal Design in Nonlinear Multiresponse Estimation: Poisson Model for Filter Feeding

Christos Hatzis and Kinley Larntz*

Department of Chemical Engineering and Materials Science and
Institute for Advanced Studies in Biological Process Technology,
University of Minnesota, St. Paul, Minnesota 55108, USA

and

Department of Applied Statistics, University of Minnesota,
St. Paul, Minnesota 55108, USA

February, 1991

* Author to whom correspondence should be addressed

Key words: Optimal design; multiresponse estimation; subset of parameters; constrained optimization; simplex algorithm; simulated annealing; protozoan filter feeding; stochastic modeling; sensitivity analysis

SUMMARY

D-optimality criteria have been applied to construct locally optimal designs for a multiresponse, nonlinear model. Simulated annealing was used to perform the needed numerical optimization calculations, as this method can locate the global optimum of a function, and can efficiently handle constraints in the independent variables. The calculated optimal designs greatly reduce variances of model parameter estimates, compared to variances from previously used empirical designs. The effect of several design variables, including the number of design points and the number of responses, on the efficiency of the design was investigated, and designs for various subsets of parameters were also calculated. New directions for the design of future experiments were suggested by this analysis.

1. Introduction

Mixed microbial communities are the rule rather than the exception in natural ecosystems, and understanding the interactions between the several microbial species has been a challenging problem for microbial ecologists and bioengineers. In our attempt to understand interactions between populations of bacteria and ciliated protozoa, which are motile microscopic organisms that feed on bacteria, we have formulated a stochastic, mechanistic model for feeding of protozoa on bacteria (Hatzis *et al.*, 1990a). Since protozoa take up the bacteria by virtually filtering them out of the suspended water, this mode of feeding is known as filter feeding. Under certain conditions the process of feeding of a single protozoan can be described as a Poisson random process, and the fraction of the overall population of protozoa that have i bacteria ingested at time t, $\eta_i(t)$, is given by

$$\eta_i(t) = (1-\theta_1)\,\delta_{i,0} + \theta_1 \int_0^\infty \frac{(\theta_2 b\tau)^i}{i!}\, e^{-\theta_2 b\tau}\, f(\theta_2)\, d\theta_2, \quad i = 0, 1, 2, \ldots \tag{1}$$

where $\delta_{i,0} = 1$ for $i = 0$ and 0 for any other i, θ_1 represents the fraction of cells in the population of protozoa that are actively feeding, θ_2 and $f(\theta_2)$ are the normalized rate of ingestion, or clearance rate, and its probability density function respectively, b is the concentration of bacteria at the beginning of the experiment, and τ is given by

$$\tau(t) = \frac{1}{P\theta_1\theta_2}(1 - e^{-P\theta_1\theta_2 t}) \tag{2}$$

with P representing the concentration of protozoa. Assuming that $f(\theta_2)$ is a gamma, Eq. (1) becomes

$$\eta_i(t) = (1-\theta_1)\,\delta_{i,0} + \theta_1 \frac{(\theta_2 b\tau)^i}{i!}(1 + \theta_3^2\theta_2 b\tau)^{-1/\theta_3^2 - i} \prod_{k=1}^{i-1}(1 + k\theta_3^2), \quad i = 0, 1, 2, \ldots \tag{3}$$

with the additional parameter θ_3 being the coefficient of variation of the distribution $f(\theta_2)$. Parameters b and P are treated as constants fixed by the experimenter.

All three parameters in the above model are essential in evaluating the

ecological role of the protozoa, and they need to be estimated from controlled experiments. Experiments are carried out in which the protozoa are offered bacteria or other particles of similar size and the fractions of the protozoan population that have 0, 1, 2, and 3 particles ingested are measured over time. These data are then used to fit the model of Eq. (3) and estimate the parameters involved. Details about the experimental and fitting procedures can be found in Hatzis *et al.* (1990b). The accuracy of the the least squares estimates of the parameters depends on the sampling schedule employed. The optimal design problem consists in determining the schedule that produces estimates of minimum variance.

Nonlinear multiresponse estimation is common in fields such as chemical kinetics (Ziegel and Gorman, 1980; Ford, Titterington, and Kitsos, 1989) and pharmacokinetics (Jacquez, 1972), and the needed methodology and theory are well developed (Bates and Watts, 1988; Seber and Wild, 1989). Design for nonlinear models originates from the work of Fisher on the serial dilution problem (see Cochran, 1973) and has been recently reviewed by Atkinson (1988) and Ford *et al.* (1989). The most commonly used design criterion for nonlinear estimation is that of D-optimality, according to which a design is selected that minimizes the generalized variance (volume of joint confidence ellipsoid) of the estimated parameters. An intrinsic difficulty of the nonlinear design problem is that the covariance structure of the estimated parameters depends on the unknown parameters themselves. The commonly used strategy is to find the optimal design for specific given values of the parameters. Such designs are known as *locally optimal* designs (Chernoff, 1953).

In this paper we use optimal design theory to construct locally optimal designs for the nonlinear multiresponse model of Eq. (3). In the next section we review briefly the criteria used for constructing these designs. The algorithms used to perform the constrained optimization calculations are described in Section 3. In Section 4, optimal designs are calculated for a certain set of conditions and compared to the empirical designs used in previous experiments. In Section 5 a sensitivity analysis for the problem is reported which shows, among others, that the efficiency of the design increases with increasing number of design points and with increasing number of responses. Finally, the problem of finding optimal designs for subsets of parameters is discussed in Section 6. We conclude in Section 7 with a discussion of the results and recommendations.

2. Locally Optimal Designs

We will now review the needed main results from optimal design theory. Consider the general nonlinear multiresponse model

$$y_{iu} = \eta_i (t_u ; \theta) + \varepsilon_{iu} \quad , \qquad i = 1, ..., M, \;\; u = 1, ..., N, \tag{4}$$

in which N sets of observations are taken on each one of the M responses. The same model can be written in terms of the N-dimensional vectors as

$$\mathbf{y}_i = \eta_i (\mathbf{t} ; \theta) + \mathbf{e}_i \quad , \qquad i = 1, ..., M,$$

or finally in matrix form as

$$Y = H + E \tag{5}$$

where,

$$Y = (\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_M),$$
$$H = (\eta_1, \eta_2, ..., \eta_M),$$

and

$$E = (\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_M) = (\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_N)^T,$$

are the $N \times M$ matrices of observations, responses, and errors, respectively. We further make the following assumptions about the error structure of the problem:

$$\mathbf{e}_i \sim \mathrm{N}(\mathbf{0}, \sigma_{ii} I_N) \quad , \qquad i = 1, ..., M \tag{6a}$$

$$\mathbf{u}_j \sim \mathrm{N}(\mathbf{0}, \Sigma) \quad , \qquad j = 1, ..., N \tag{6b}$$

which are equivalent to

$$E \sim \mathrm{N}(\mathbf{0}, \Sigma \otimes I_N) . \tag{6c}$$

The first assumption means that errors on each response are the same for all experimental runs and that errors on measurements from different experiments are uncorrelated. This implies that the error vectors $\mathbf{u}_j$ (error vector for all responses from run j) are independent and according to the second assumption follow a multivariate normal distribution with mean $\mathbf{0}$ and variance-covariance matrix Σ.

(The symbol $\otimes$ is used for the Kronecker or direct product of two matrices.)

The design problem consists of selecting N points $\{t_u, u = 1, ..., N\}$ for measurement of the M responses. In realistic situations there will be constraints on the points, such as the t_u must be within certain limits and the t_u must differ from each other by at least a given amount. We account for these constraints in our optimal design computations. Box and Lucas (1959) and then Draper and Hunter (1966) applied the D-optimality criterion to a linearized form of the model of Eq. (5). If we denote by X_j the $N \times P$ matrix of derivatives of the response vector η_j with respect to the P-dimensional parameter vector θ calculated at a specified point θ^o,

$$X_j^T = \left. \frac{\partial \eta_j (t ; \theta)}{\partial \theta} \right|_{\theta = \theta^o}, \tag{7}$$

then the D-optimality criterion for the multiresponse case is

$$\max_{t_u} \{ \det (\Delta) \}, \tag{8}$$

that is, the optimal design, $\{t_u, u = 1, ..., N\}$, is the one that maximizes the determinant of the information matrix Δ, where

$$\Delta = V^T \, (\Sigma^{-1} \otimes I_N) \, V \tag{9}$$

is a $P \times P$ matrix and $V(\theta^o)$ is the $NM \times P$ matrix of response derivatives,

$$V^T = (X_1^T \ldots X_M^T) = \left. \frac{\partial \, \mathrm{vec}H}{\partial \theta} \right|_{\theta = \theta^o}. \tag{10}$$

(The symbol vec$H$ is the vector of matrix H (see e.g. Graham, 1981).) Since the information matrix Δ depends on θ^o and the unknown variance-covariance matrix Σ, the resulting design is a locally optimal design. In practice, initial estimates for θ and Σ are obtained from past data and scientific knowledge about the problem.

There are situations in which only a subset of parameters is of primary interest. If only the first Q of the P parameters need to be estimated precisely, Box (1971) and Hill and Hunter (1973) have shown that the criterion applicable to these

cases is

$$\max_{t_u} \{ \det (\Delta_s) \} \tag{11}$$

with the information matrix now being

$$\Delta_s = V_{11} - V_{12}\, V_{22}^{-1} V_{21} \tag{12}$$

where V_{11}, V_{12}, V_{21}, and V_{22} are respectively the $Q \times Q$, $Q \times P-Q$, $P-Q \times Q$, and $P-Q \times P-Q$ submatrices of Δ. It can be shown (see e.g. Graybill, 1983) that if V_{22} is nonsingular then

$$\det (\Delta_s) = \frac{\det (\Delta)}{\det (V_{22})} , \tag{13}$$

so that only an additional calculation of the determinant of V_{22} is needed for the application of the subset design criterion.

We conclude this section with the inference problem. Kang and Bates (1990) showed that the variance-covariance matrix of the maximum likelihood estimator $\hat{\theta}$ is

$$\mathrm{cov}(\hat{\theta}) = \Delta^{-1} = \{ V^T \, (\Sigma^{-1} \otimes I_N) \, V \}^{-1}. \tag{14}$$

Also $\hat{\theta}$ is asymptotically normally distributed under mild regularity conditions. These results are useful for constructing confidence regions for the estimates of θ, and also for comparing efficiencies of alternative designs.

3. Optimization Algorithms

Several iterative optimization algorithms have been developed and used for calculating *D*-optimal designs. Cook and Nachtsheim (1980) and Johnson and Nachtsheim (1983) reviewed deterministic algorithms available for constructing exact *D*-optimal designs. The main problem in such calculations is that the functions to be maximized have multiple maxima and these algorithms cannot guarantee locating the global maximum. A promising new class of algorithms for global optimization is the stochastic search algorithms and, in particular, the

simulated annealing algorithm (Kirkpatrick, Gelatt, and Vecchi, 1983). If certain conditions for the parameters are met, this algorithm is guaranteed to converge to the global optimum (Laarhoven and Aarts, 1987). Haines (1987) used the original version of the algorithm to construct exact optimal designs for linear models, and Bohachevsky, Johnson, and Stein (1986) used a modified version to construct a *D*-optimal design for a nonlinear model.

To calculate optimal designs we found that a combination of a simple downhill algorithm, such as the simplex algorithm of Nelder and Mead (1965), and the generalized simulated annealing (GSA) algorithm of Bohachevsky *et al.* (see also Kalivas, Roberts and Sutter, 1989) gave the best results. The simplex algorithm was started from a random point in the state space and the calculated optimum was used as an initial estimate for GSA. To adjust the controlling parameters for optimal performance of the GSA algorithm, we had to perform several exploratory runs. The values of the parameters used in the GSA optimizations reported in this study were $5 \leq \beta \leq 40$ with most frequent value 20, $5 \leq \Delta r \leq 10$, and $g = 1$, which correspond to a 30-60% acceptance of unfavorable steps. The GSA algorithm was stopped and convergence to the global optimum was assumed when at least 50 steps were taken without an acceptance. A maximum of 5000 function evaluations was imposed in order to avoid wandering of the algorithm in nonconvergent situations.

Both algorithms were adapted to handle constraints in the independent variables. In the simplex algorithm, the objective function was set equal to an arbitrarily large value (for the minimization problem) whenever the variables did not satisfy the imposed constraints. GSA is inherently more efficient in handling constraints since trial vectors that fall out of the bounds of the feasible region are rejected right after they are generated and thus do not contribute to the computational cost. In most of the cases tested GSA gave the best optimum result, although in 8 out of 55 runs the simplex algorithm located a better maximum when started from several randomly chosen initial points. The computational effort for the simplex algorithm increased with almost the square of the number of dimensions of the optimization problem, whereas for GSA the increase is approximately linear (see Fig. 1). In general, for low-dimensional problems it takes the simplex fewer function evaluations to converge than the GSA algorithm, however GSA can be equally or even more efficient at higher dimensions. The derivatives of Eq. (7) were calculated analytically, and the algorithms were written in

FORTRAN and run on a Cray-XMP supercomputer. On this computer it takes about 16 sec of CPU time for 5000 iterations of the GSA algorithm, but usually convergent trials did not take more than 2-3 sec.

4. Example

We applied the theory outlined in Section 2 to find the optimal design for the protozoan feeding experiment described in the Introduction. In previous experiments we used an empirical sampling schedule to make observations on four responses, $\mathbf{y}_i = (y_{0i}, y_{1i}, y_{2i}, y_{3i})$, with the sampling times chosen so that they span almost all the range of interest and still meet the imposed experimental constraints. Such schedules were valuable in assessing model appropriateness during the initial stages of the model building process. Since all three parameters are involved in all responses of the model of Eq. (3), we used information from all four measured responses. For five different data sets, iteratively reweighted least squares (Green, 1984) was used to fit the model for $i = 0, 1, 2$, and 3 to the sets of observations and give robust estimated parameters $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3)$. The variance-covariance matrix $\hat{\Sigma}$ can be obtained from the residuals of the fit as

$$\hat{\Omega} = \hat{\Sigma} \otimes I_N = \frac{\hat{E}^T \hat{E}}{N - P}. \tag{15}$$

The averages of the estimates obtained from the five different data sets are given in Table 1.

To evaluate the previously used empirical designs we calculated the corresponding locally optimal designs and compared the efficiencies of the two designs. Three different values of parameter b were considered, and all designs involved seven sampling points. The constraints imposed on the sampling times were: (i) the first sample should be taken at least 10 sec after the beginning of the experiment ($t_1 \geq 10$), (ii) a minimum interval of 30 sec should separate consecutive samples ($t_u - t_{u-1} \geq 30$), and (iii) the maximum duration of the experiment should not exceed 1000 sec ($t_N \leq 1000$). The first two constraints are results of the sampling procedure and the preparations required before a sample is taken, and the third one ensures that changes in the cell populations due to growth will be minimal during

the course of the experiment.

The resulting designs are shown in Figure 2, where the four responses of the model are also drawn. The optimal designs for the first two cases were significantly different than the empirical ones, with the optimal designs placing greater support (more sampling points) at later times, whereas the two designs were very similar for the last case. To compare the efficiencies and the effect that these designs have on the estimated parameters, we computed the standard deviations of the parameter estimates (see Eq. (14)) and also a measure of the relative efficiencies of the designs, $(|\Delta_e| / |\Delta_o|)^{1/2}$, where $|\Delta_e|$ is the criterion value for the empirical design and $|\Delta_o|$ is the same value for the corresponding optimal design. The results from these computations are given in Table 2. Since the volume of the asymptotic confidence ellipsoid for the estimates of $(\theta_1, \theta_2, \theta_3)$ is inversely proportional to the square root of $|\Delta|$ (see e.g. Seber and Wild, 1989), the relative efficiency represents the fraction by which the generalized variance of the estimates is reduced when parameters are estimated according to the optimal schedule. Relative efficiency (RE) of 0.476, for instance, implies that the volume of the confidence ellipsoid obtained from the optimal design is 47.6% of the volume of the corresponding ellipsoid obtained from the empirical design, or that the average confidence interval of any of the parameters will be $(0.476)^{1/3} = 0.781$ or 78.1% of the corresponding empirical one. These results demonstrate that the optimal designs do not always coincide with the intuitively appealing ones, and in addition, that we can achieve over a 50% reduction in the variance of the estimated parameters by designing the experiments appropriately.

5. Effects of Design Variables

The previous example showed that an experimenter has much to gain from a proper experimental design. However, the design itself depends on a set of variables, which include the number of design or support points, the number of responses used, the error variance-covariance structure, and the set of imposed constraints. Determination of the effects of these variables on the final design might be important in pointing out possible limitations of the experimental protocols in use, or even in indicating potential ways of modifying the protocols for increased efficiency. For instance, two designs, one with four support points and data on four

responses and another one with seven points and data on three responses, can be equally efficient. The former design however is a more economical one since it requires less sampling effort and more analysis.

Motivated by these considerations, we carried out a systematic investigation of the effects of the aforementioned design variables on the efficiency of the resulting optimal design. The design of the previous section (at $b = 1x10^6$) was used as the standard design, and the design variables were varied one at a time to determine how each variable affects the optimal design. Since experimental effort increases with the number of support points, we calculated optimal designs for variable numbers of support points and compared their efficiencies. The results are given in Table 3 and the corresponding designs are shown in Figure 3. Figure 3a reveals that the support points are grouped into two distinct clusters, with the first cluster located close to the regions of steep change, or where the derivatives of the responses are large, and the second cluster located close to the asymptotes of the responses (see also Figure 2b). Efficiencies were calculated relative to the best design (10 support points), and these numbers are plotted in Figure 3b as relative reduction in the generalized variance of the estimates of θ. As one could have anticipated, the design with just one support point still gives enough information for estimating the three parameters, since it involves data from four responses. However, the resulting estimates are very inaccurate. The accuracy is improved very rapidly with increasing number of support points, but the improvement becomes proportionately less as the designs become larger. This analysis indicates that the optimal size for such a design is around six or seven support points, since for larger designs the higher cost of experimentation is not compensated for by the gain in quality of the estimates.

Since all parameters appear in each of the responses, the number of responses is expected to affect the efficiency of the design. The resulting designs for various numbers of responses are given in Table 4 and also plotted in Figure 4 together with the corresponding responses. The RE of the one-response design is very poor, and considerable improvement can be achieved by including data from increasing number of responses. Parameter θ_1 can be estimated with adequate accuracy from just the first response and small improvement is gained if more responses are also considered. On the contrary, the standard deviation of the other two parameter estimates decreases drastically when any additional response is included in the

design. These results suggest that the potential advantage from including larger number of responses in the analysis of the data should be considered in future experiments. Besides its practical importance, the above analysis can be used for locating the information-rich areas for parameter estimation for each set of responses, by simply looking at the designs of Figure 4.

The several constraints imposed by the experimental procedure might alter the resulting design. The effect of the minimum interval between consecutive samples was examined, and the resulting designs are shown in Table 5. It appears that reducing the minimum sampling interval from 30 to 20 or even 10 sec does not have an apparent effect on the efficiency of the design. Thus, the more convenient sampling interval of 30 sec can be used without great loss in efficiency.

Finally, we investigated the effect of the variance-covariance structure of the error for testing the robustness of the resulting locally optimal designs. If instead of the estimated variance-covariance matrix $\hat{\Sigma}$ we use the matrix $\hat{\sigma}^2 I$, with $\hat{\sigma}^2$ being the average error variance, the resulting design is different and, in fact, considerably less efficient (see Table 6). This indicates that the local designs are not very robust with respect to the values of the unknown parameters.

6. Estimation of Subsets of Parameters

In the designs of the previous sections we assumed that all parameters need to be estimated with the same accuracy. However, there might be cases in which only a subset of the parameters is needed. For example, parameters θ_1 and θ_2 can be determined directly from alternative measurements without making use of the model of Eq. (3) (Hatzis *et al.*, 1990b). In this case, experiments could be designed focussing mainly on the optimal estimation of the third parameter θ_3.

We constructed optimal designs for the estimation of all posible subsets of parameters, and these designs are shown in Figure 5. It is evident from this figure that designs can change considerably depending on the parameters that need to be estimated. In other words, different regions of the design space contain information on different parameters (see Figure 5b). The designs based on the overall criterion or on larger subsets of parameters seem to be compromises between the one-

parameter subset designs.

The above designs are also given in Table 7. For a more direct comparison of the effect that the various criteria have on the accuracy of the estimates, we computed the percent change in the standard deviations of the estimates relative to the complete design. We can see, for instance, that if we design mainly for parameter θ_1 its estimate can be improved by 12.5%, but the error in the other two parameters becomes prohibitely large so that such a design cannot be of practical use. If we design for better estimation of the second parameter, θ_2, we gain 16% in accuracy, but the error in the estimate of θ_3 increases by 30%. Overall, it appears that the designs are more robust to the first parameter, in the sense that this parameter can be estimated accurately irrespective of the design used. On the other hand, there seems to be no significant improvement in the estimates of the other two parameters by using the subset criteria.

7. Discussion

In this paper, we applied the existing theory of *D*-optimal design to a nonlinear, multiresponse model which describes the behavior of a biological system. For the calculation of the optimal designs we used the stochastic search technique of simulated annealing (Bohachevsky *et al.*, 1986) for determining the global optimum of a continuous function under multiple constraints. With this technique, the location of the global maximum of the criterion function is almost certain, so that the resulting designs are truly optimal and not near-optimal designs.

One can argue that locally optimal designs are not very useful in practice since parameters θ are not known at the stage of the design. However, as Ford *et al.* (1989) comment, such designs are of interest since they provide the reference point for sequential or more robust designs. In fact, in many instances locally optimal designs can be reasonably stable over ranges of θ. As an alternative to using a best guess of the parameter values, one can follow the Bayesian approach and use all the available prior information on the unknown parameters. The uncertainty in parameters θ can be represented by a prior distribution, and the criterion value has to be averaged, in some way, over this prior distribution. The resulting designs are *exact* optimal designs, and are more robust to different initial parameter values

(Atkinson, 1988; Chaloner and Larntz, 1989). Chaloner and Larntz (1989) have developed appropriate criteria for calculating exact optimal designs, and they have demonstrated their methodology for a logistic regression model. However, application of this type of design criteria for calculation of exact optimal designs for the problem in hand will probably lead to formidable calculations, since numerical integrations in a $[P + M(M + 1)/2]$-dimensional space need to be carried out if prior information on both θ and Σ are to be used.

For the particular example that we analyzed, the D-optimal designs were found to be more efficient than the empirical designs based on uniform spacing. Sensitivity analysis on the calculated locally optimal designs revealed that both the number of design points and the number of responses increase the efficiency of the design. On the other hand, the experimental constraints did not appear to affect the efficiency of the calculated designs. By replacing the estimated variance-covariance matrix with a diagonal matrix, we showed that the locally optimal designs are sensitive to the estimates of the unknown parameters, and thus are not very robust. In contrast to what we had hoped, the designs based on subset criteria did not result in considerable improvement in the estimation of selected parameters.

Overall, the sensitivity analysis on the optimal designs suggested that it is advantageous for the experimenter to put more effort into the analysis of the data than into sampling, since substantial improvement can be achieved by including data on larger number of responses. However, due to nature of the error inherent in the measurement techniques, it becomes increasingly difficult to retrieve error-free data on larger numbers of responses (Fredrickson *et al.*, 1991). Consequently, the design of an efficient experimental schedule would have to be a compromise between the cost of taking larger numbers of samples, and the accuracy of measurements on increasing numbers of responses. Work toward this latter direction is under way (Fredrickson *et al.*, 1991).

ACKNOWLEDGEMENTS

C. Hatzis wishes to thank Professors A. G. Fredrickson and F. Srienc for their helpful comments and suggestions. This research was supported in part by National Science Foundation grants, BCS-8619399-02 and DMS-8706754.

REFERENCES

Atkinson, A. C. (1988). Recent developments in the methods of optimum and related experimental designs. *International Statistical Review* **56,** 99-115.

Bates, D. M. and Watts, D. G. (1988). *Nonlinear Regression Analysis and its Applications*. New York: Wiley.

Bohachevsky, I. O., Johnson, M. E., and Stein, M. L. (1986). Generalized simulated annealing for function optimization. *Technometrics* **28,** 209-217.

Box, G. E. P. and Lucas, H. L. (1959). Design of experiments in nonlinear situations. *Biometrika* **46,** 77-90.

Box, M. J. (1971). An experimental design criterion for precise estimation of a subset of the parameters in a nonlinear model. *Biometrika* **58,** 149-153.

Chaloner, K. and Larntz, K. (1989). Optimal Bayesian design applied to logistic regression experiments. *Journal of Statistical Planning and Inference* **21,** 191-208.

Chernoff, H. (1953). Locally optimal designs for estimating parameters. *Annals of Mathematical Statistics* **24,** 586-602.

Cochran, W. G. (1973). Experiments for nonlinear functions. *Journal of the American Statistical Association* **68,** 771-781.

Cook, R. D. and Nachtsheim, C. J. (1980). A comparison of algorithms for constructing exact D-oprimal designs. *Technometrics* **22,** 315-324.

Draper, N. R. and Hunter, W. G. (1966). Design of experiments for parameter estimation in multiresponse situations. *Biometrika* **53,** 525-533.

Ford, I., Titterington, D. M., and Kitsos, C. P. (1989). Recent advances in nonlinear experimental design. *Technometrics* **31,** 49-60.

Fredrickson, A. G., Srienc, F., Hatzis, C., and Sweeney, P. J. (1991). Statistics of phagocytosis as determined by flow cytometry. Submitted to *Cytometry*.

Graham, A. (1981). *Kronecker Products and Matrix Calculus with Applications*. Chichester, England: Ellis Horwood.

Graybill, F. A. (1983). *Matrices with Applications in Statistics*, 2nd ed. Belmont, Ca.: Wadsworth.

Green, F. J. (1984). Iteratively reweighted least squares for maximum likelihood estimation, and some robust and resistant alternatives (with discussion). *Journal of the Royal Statistical Society, Series B* **46,** 149-192.

Haines, L. M. (1987). The application of the annealing algorithm to the construction

of exact optimal designs for linear-regression models. *Technometrics* **29,** 439-447.

Hatzis, C., Sweeney, P. J., Srienc, F., and Fredrickson, A. G. (1990a). A discrete, stochastic model for microbial filter feeding: a model for feeding of ciliated protists on spatially uniform, nondepletable suspensions. *Mathematical Biosciences* **102,** 127-181.

Hatzis, C., Sweeney, P. J., Srienc, F., and Fredrickson, A. G. (1990b). Distribution of rates of particle uptake from dilute suspensions by the ciliate *Tetrahymena pyriformis*. Submitted to *Journal of General Microbiology*.

Hill, W. J. and Hunter, W. G. (1974). Design of experiments for subsets of parameters. *Technometrics* **16,** 425-434.

Jacquez, J. A. (1972). *Compartmental Analysis in Biology and Medicine.* New York: Elsevier.

Johnson, M. E. and Nachtsheim, C. J. (1983). Some guidelines for constructing exact *D*-optimal designs on convex design spaces. *Technometrics* **25,** 271-277.

Kalivas, J. H., Roberts, N., and Sutter, J. M. (1989). Global optimization by simulated annealing with wavelength selection for ultraviolet-visible spectrophotometry. *Analytical Chemistry* **61,** 2024-2030.

Kang, G. and Bates, D. M. (1990). Approximate inferences in multiresponse regression analysis. *Biometrika* **77,** 321-331.

Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. (1983). Optimization by simulated annealing. *Science* **220,** 671-680.

Laarhoven, P. J. M. van and Aarts, E. H. L. (1987). *Simulated Annealing: Theory and Applications*. Dordrecht, Holland: Reidel.

Nelder, J. A. and Mead, R. (1965). A simplex method for function minimization. *Computer Journal* **7,** 308-313.

Seber, G. A. F. and Wild, C. J. (1989). *Nonlinear Regression*. New York: Wiley.

Ziegel, E. R. and Gorman, J. W. (1980). Kinetic modelling with multiresponse data. *Technometrics* **22,** 139-151.

FIGURE CAPTIONS

Fig. 1. Comparison of performance of generalized simulated annealing (GSA) and simplex algorithms. The number of function evaluations (fe) is plotted as a function of the dimension (d) of the optimization problem. For GSA, fe ~ $d^{0.94}$, whereas for the simplex fe ~ $d^{1.70}$.

Fig. 2. Comparison of calculated locally optimal designs with corresponding empirical ones: (o) optimal points, (●) empirical points. The four predicted responses are also plotted. Plots are for three values of parameter b: (a) 3×10^5, (b) 1×10^6, and (c) 3×10^6.

Fig. 3. Effect of number of design points: (a) layout of locally optimal designs calculated for variable number of support points, (b) generalized variance of parameter estimates relative to the maximum-size design, for designs of increasing size. (In all designs, number of responses, $M = 4$, and $b = 1\times10^6$.)

Fig. 4. Effect of number of responses: designs resulted from experiments with one to four responses, plotted with the corresponding predicted responses. (Number of design points, $N = 7$, and $b = 1\times10^6$.)

Fig. 5. Designs for subsets of parameters: (a) designs based on two-parameter subsets and (b) designs based on single-parameter subsets. In both cases, the complete design seems to be a combination of the subset designs ($N = 7$, $M = 4$, and $b = 1\times10^6$).

Table 1

Robust estimates of model parameters from previous experiments

$(\hat{\theta}_1,\ \hat{\theta}_2,\ \hat{\theta}_3)$	*Variance-Covariance Matrix* $\hat{\Sigma}^{\S}$			
	$2.00\text{x}10^{-2}$			
$(0.869,\ 2.24\text{x}10^{-8},\ 0.404)$	-0.544	$2.06\text{x}10^{-2}$		
	-0.387	0.550	$8.35\text{x}10^{-3}$	
	-0.153	-0.148	0.213	$5.09\text{x}10^{-3}$

(§) The diagonal elements of matrix $\hat{\Sigma}$ are the standard deviations $\hat{\sigma}_i$ and the off-diagonal elements are the correlation coefficients $\hat{\rho}_{ij}$

Table 2

Comparison of optimal and empirical designs

b	N	$\{t_u\}$ (sec)	$s(\hat{\theta}_1)$	$s(\hat{\theta}_2)$	$s(\hat{\theta}_3)$	$ln\|\Delta(t_u)\|$	RE
Empirical	*designs*						
$3x10^5$	7	20 120 240 360 480 600 720	0.013	$3.778x10^{-10}$	0.031	59.991	0.476
$1x10^6$	7	30 60 90 120 180 240 360	0.010	$3.864x10^{-10}$	0.027	60.381	0.658
$3x10^6$	7	30 60 90 120 180 240 360	0.007	$4.616x10^{-10}$	0.027	60.676	0.924
Optimal	*designs*						
$3x10^5$	7	548 582 616 648 937 970 1000	0.008	$2.975x10^{-10}$	0.023	61.476	-
$1x10^6$	7	145 175 205 235 405 443 473	0.008	$3.904x10^{-10}$	0.021	61.219	-
$3x10^6$	7	21 53 84 116 150 186 363	0.007	$4.344x10^{-10}$	0.025	60.835	-

Table 3

Effect of number of design points

N	$\{t_u\}$ (sec)										$ln\|\Delta(t_u)\|$	RE
1	175										54.980	0.028
2	173	406									57.553	0.101
3	165	196	419								58.769	0.185
4	161	191	383	442							59.610	0.282
5	151	182	212	402	458						60.280	0.394
6	154	184	218	308	404	472					60.779	0.506
7	145	175	205	235	406	443	473				61.219	0.630
8	137	168	199	230	265	377	419	493			61.594	0.760
9	121	151	181	212	257	371	405	475	505		61.897	0.884
10	135	165	195	225	295	395	443	496	526	556	62.143	1.000

Table 4

Effect of number of responses

M	$\{t_u\}$ (sec)							$s(\hat{\theta}_1)$	$s(\hat{\theta}_2)$	$s(\hat{\theta}_3)$	$ln\|\Delta(t_u)\|$	RE
1	23	53	120	150	912	970	1000	0.012	2.012×10^{-9}	0.143	54.308	0.032
2	51	82	171	201	233	734	1000	0.011	1.087×10^{-9}	0.073	56.949	0.118
3	95	125	156	187	295	369	418	0.009	5.713×10^{-10}	0.032	59.574	0.439
4	145	175	205	235	405	443	473	0.008	3.904×10^{-10}	0.021	61.219	1.000

Table 5

Effect of minimal sampling interval

τ (sec)	$\{t_u\}$ (sec)							$ln\|\Delta(t_u)\|$	RE
10	157	169	181	191	380	393	404	61.297	1.000
20	148	168	189	210	377	401	422	61.270	0.987
30	145	175	205	235	405	443	473	61.219	0.962

Table 6

Effect of variance-covariance structure of error

Variance-Covariance Matrix[§]	$\{t_u\}$ *(sec)*							$ln\|\Delta(t_u)\|$	RE
$\hat{\Sigma}$	145	175	205	235	405	443	473	61.219	1.000
$\hat{\sigma}^2 I$	84	115	145	175	206	382	487	59.938	0.527

(§) $\hat{\Sigma}$ is the estimated error variance-covariance matrix
I is the identity matrix and $\hat{\sigma}^2$ is the average estimated variance

Table 7

Designs for subsets of parameters

Design Criterion	*t (sec)*							$s(\hat{\theta}_1)$	$s(\hat{\theta}_2)$	$s(\hat{\theta}_3)$	$ln\lvert\Delta_s\rvert$	% Change in $s(\hat{\theta}_1)$	% Change in $s(\hat{\theta}_2)$	% Change in $s(\hat{\theta}_3)$
$\Delta(\theta_1, \theta_2, \theta_3)$	145	175	205	235	405	443	473	0.008	3.904×10^{-10}	0.021	61.219	0.0	0.0	0.0
$\Delta_s(\theta_1, \theta_2)$	165	196	226	256	289	398	520	0.008	3.662×10^{-10}	0.024	53.145	0.0	-6.2	14.3
$\Delta_s(\theta_1, \theta_3)$	132	162	192	491	524	558	600	0.008	4.900×10^{-10}	0.019	17.919	0.0	25.5	-9.5
$\Delta_s(\theta_2, \theta_3)$	126	156	186	216	355	385	415	0.008	3.866×10^{-10}	0.021	51.192	0.0	-1.0	0.0
$\Delta_s(\theta_1)$	383	720	761	791	922	970	1000	0.007	1.685×10^{-9}	0.045	10.007	-12.5	331.6	114.3
$\Delta_s(\theta_2)$	51	148	178	208	239	270	302	0.009	3.279×10^{-10}	0.027	43.677	12.5	-16.0	28.6
$\Delta_s(\theta_3)$	124	155	185	478	509	545	581	0.008	5.001×10^{-10}	0.019	7.898	0.0	28.1	-9.5

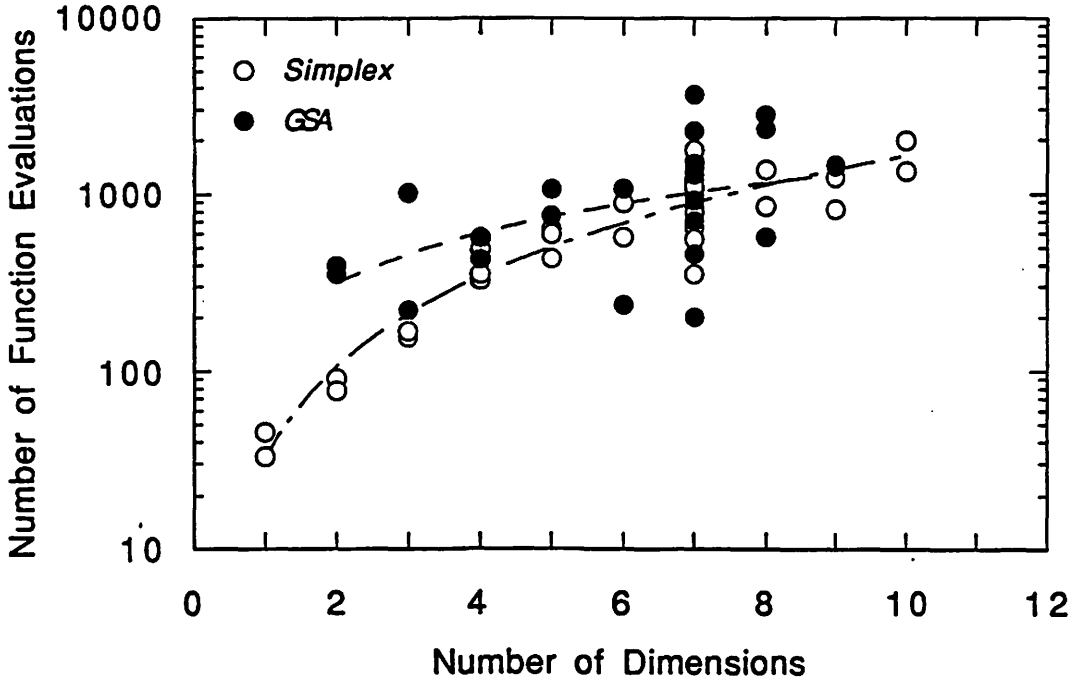

10000
1000
100
10
Number of Function Evaluations
Simplex
GSA
0
2
4
6
8
10
12
Number of Dimensions

Figure 1

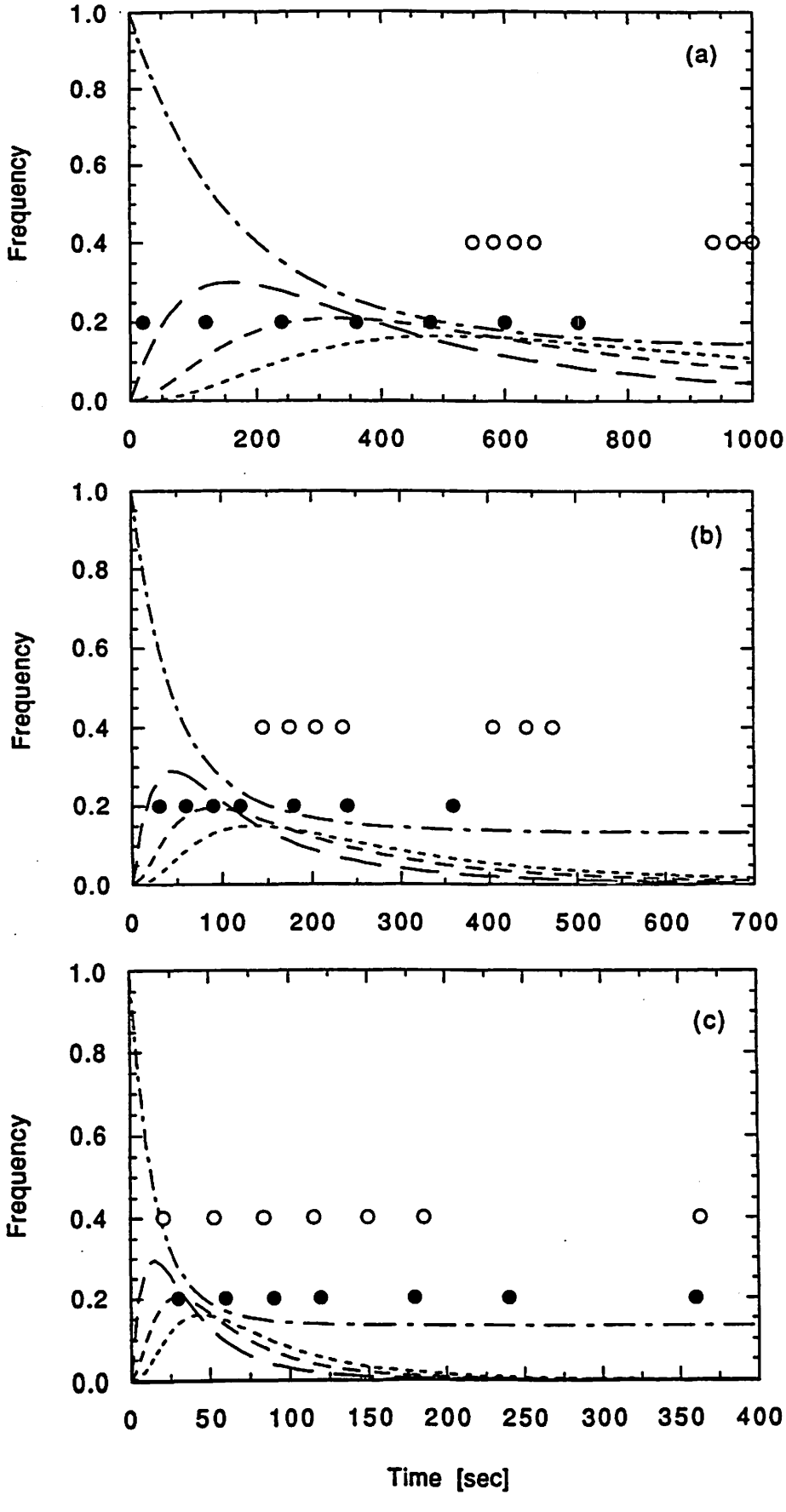
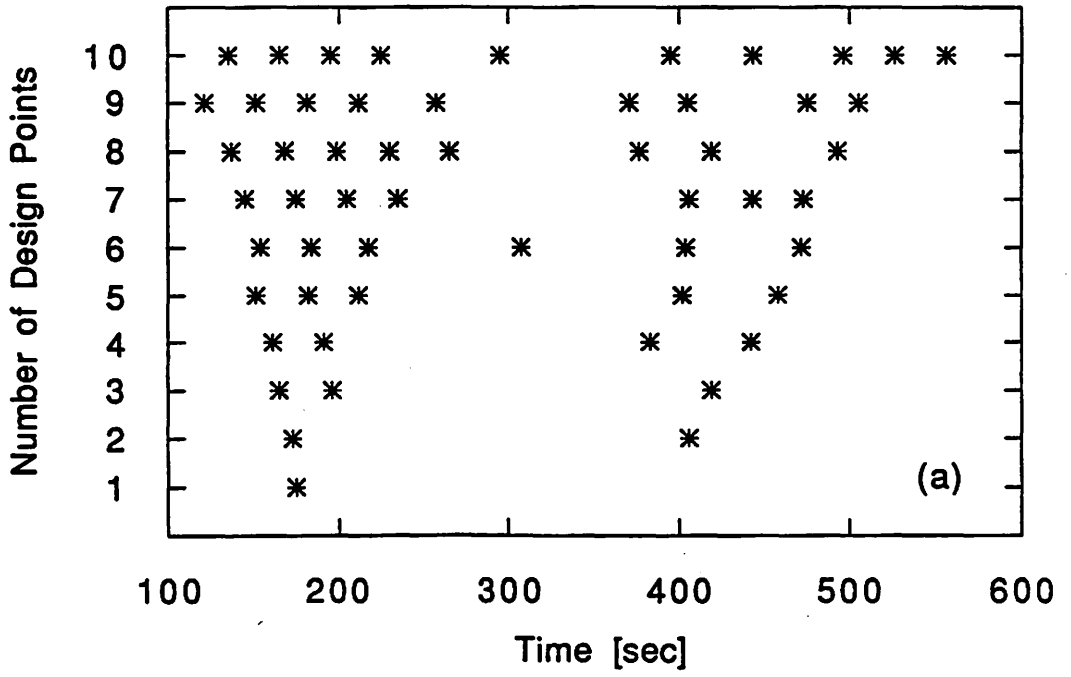

(a)
(b)
(c)
Frequency
Time [sec]
1.0
0.8
0.6
0.4
0.2
0.0
0
200
400
600
800
1000
100
300
500
700
50
150
250
350

Figure 2


Number of Design Points
10
9
8
7
6
5
4
3
2
1
(a)
100
200
300
400
500
600
Time [sec]

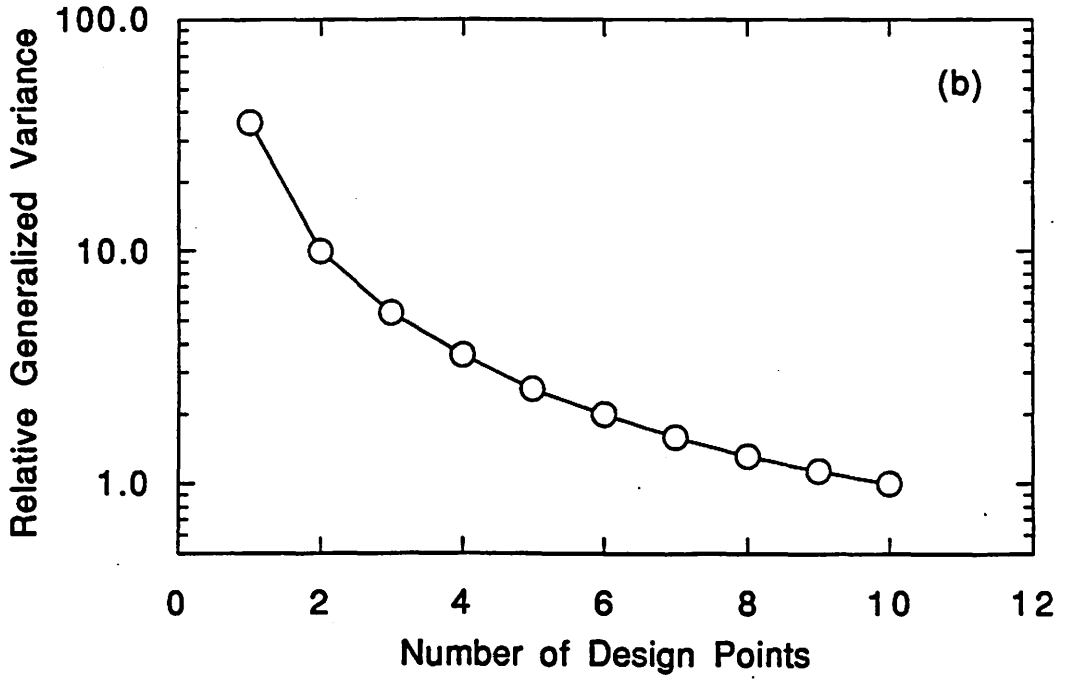

Relative Generalized Variance
100.0
10.0
1.0
(b)
0
2
4
6
8
10
12
Number of Design Points

Figure 3

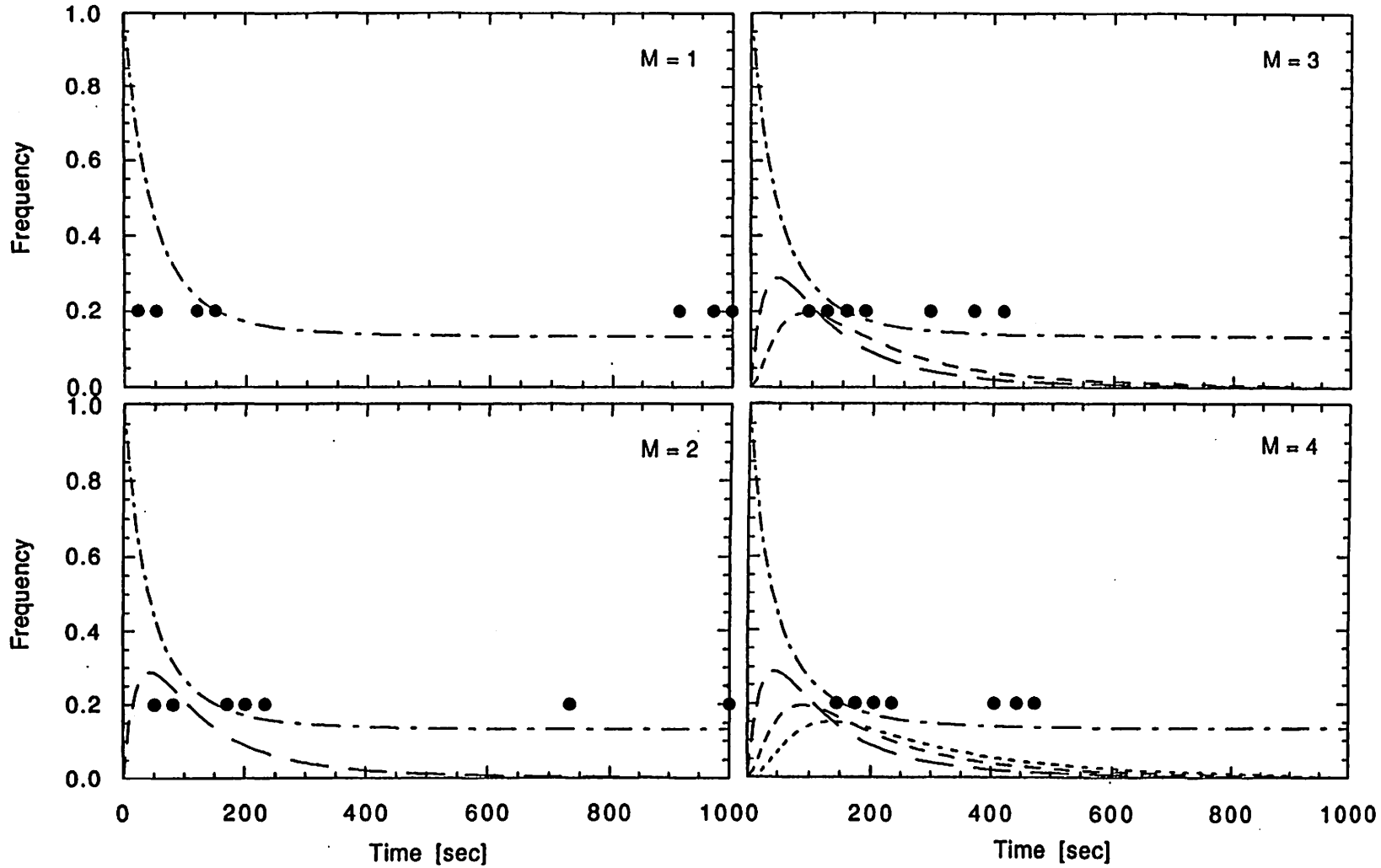

M = 1
M = 3
M = 2
M = 4
Frequency
Frequency
1.0
0.8
0.6
0.4
0.2
0.0
0
200
400
600
800
1000
Time [sec]
Time [sec]

Figure 4

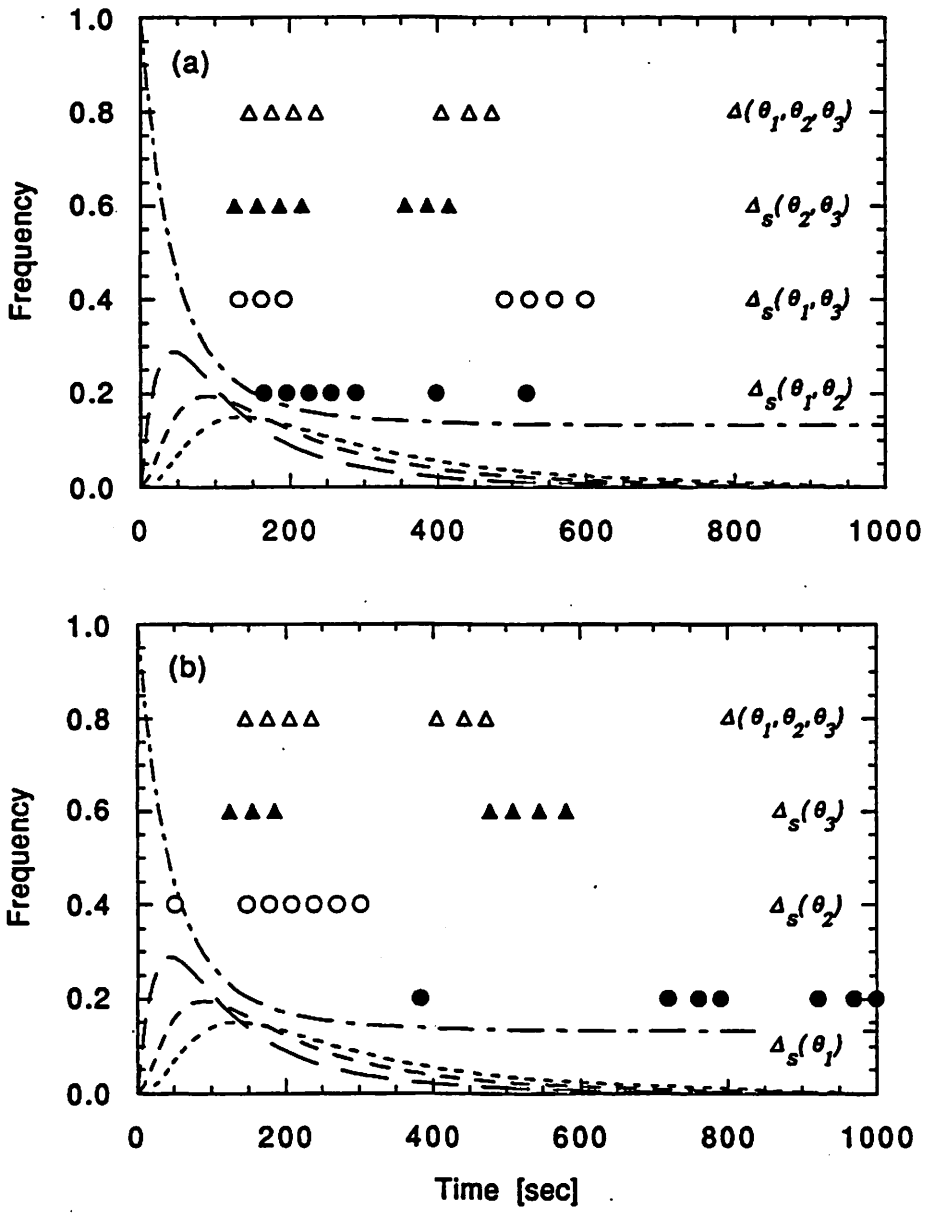

(a)
1.0
0.8
0.6
0.4
0.2
0.0
Frequency
0
200
400
600
800
1000
$\Delta(\theta_1,\theta_2,\theta_3)$
$\Delta_s(\theta_2,\theta_3)$
$\Delta_s(\theta_1,\theta_3)$
$\Delta_s(\theta_1,\theta_2)$
(b)
$\Delta(\theta_1,\theta_2,\theta_3)$
$\Delta_s(\theta_3)$
$\Delta_s(\theta_2)$
$\Delta_s(\theta_1)$
Time [sec]

Figure 5